

---

# Integrating Quantum and Molecular Mechanics

---

**ROBERT W. HARRISON**

*Kimmel Cancer Center, Department of Microbiology, Thomas Jefferson University, 233 South 10th Street, Philadelphia, Pennsylvania 19107*

*Received 6 March 1998; accepted 28 June 1999*

---

**ABSTRACT:** A computer algorithm is developed for integrating density functional quantum mechanics into a molecular mechanics program. The computationally infeasible aspects of the standard LCAO-MO approach (the iterative calculation of eigenvectors and the requirement of orthogonal expansions for the orbitals) are replaced with an efficient use of optimization via the trace theorem of linear algebra. The construction of a basis is also described for expanding the electron density that transforms with the molecular geometry. The combination of the trace method and the basis allow the solution for one configuration of atoms and electrons to be tracked over a wide range of internal conformations. The approach is readily adaptable to being used in the context of an imposed classical field that allows it to be used on part of a macromolecular complex. The initial implementation in the program AMMP is described.  
© 1999 John Wiley & Sons, Inc. J Comput Chem 20: 1618–1633, 1999

**Keywords:** quantum mechanics; molecular mechanics; trace theorem; linear algebra

---

## Introduction

**M**olecular mechanics is a classical approximation to the inherently quantum problem of describing chemistry. Despite this limitation molecular mechanics is often successful, and it is

especially useful for large problems like modeling proteins and nucleic acids. For example, molecular mechanics can be used to describe and predict the interactions between an enzyme and substrate or between enzymes and inhibitors.<sup>1,2</sup> However, there are problems for which molecular mechanics is not well suited. These are problems like “transition states,” where the molecules are severely distorted from equilibrium, and in the process of making or breaking bonds. In these cases it would be advantageous to use even an approximate quantum mechanical treatment. Quantum mechanics provides

*Correspondence to:* R. W. Harrison; e-mail: harrison@asterix.jci.tju.edu

Contract/grant sponsor: United States Public Health Service; contract/grant number: AI41380

a good basis for the choice of molecular mechanics potentials.

A number of studies of the potential utility of quantum treatment of parts of molecules in the context of a larger classical system are described in the literature.<sup>3-5</sup> In general, these methods adapt existing computer codes and algorithms to treat part of a chemical system with eigenvector expansion LCAO-MO Hartree-Fock calculations. Both *ab initio* and semiempirical methods were used. There are two major difficulties in these approaches: maintaining orthogonality when coordinates are changed and coupling the quantum and classical mechanical systems. Most of the computational difficulties with these approaches stem from the choice of the algorithm for solving the Hartree-Fock equations.

In this article a different approach is described that combines quantum mechanics for selected atoms with molecular mechanics for the parts of the problem that molecular mechanics treats well. A trace-based density functional theory is developed and implemented with a simple basis that is keyed to the molecular geometry. The electron density is parameterized so that it moves smoothly with changes in the internal geometry of the molecule. A self-consistent wave function found for one molecular conformation is smoothly deformed as the conformation changes, resulting in either a good starting point for resolving the wave function in the new conformation or a sufficiently good approximation by itself. A simple basis set is used to study the computational process and to demonstrate its feasibility. Obviously, more complete basis sets will result in higher accuracy, but only if the basic approach is feasible. The basis set is essentially an extension of floating spherical Gaussians (FSGO)<sup>6</sup> and Gaussian lobe orbitals.<sup>7</sup> While these oversimplified approaches are necessarily less accurate than a highly sophisticated basis set, they can be surprisingly good at predicting energetic trends and molecular geometries. Frost and Rouse<sup>6</sup> produced bond length errors of about 0.03 Å, which is comparable to the errors in molecular mechanics. It is designed so that it transforms with the internal coordinates of the molecule. Once the solution is found in one conformation, it can be readily tracked as the molecule moves, resulting in overall efficiency because the expensive minimization calculations do not need to be repeated. The use of the trace is critical to coupling the quantum solution to the classical world that surrounds it.<sup>8</sup> In order to include the

interactions among the classical parts of the problem, it is only necessary to include an approximation for the elements on the diagonal of the Hartree-Fock matrices. For the electronic terms this is simply the same electron-nuclear interaction integral used for the quantum calculations but instead using the coordinates of the classical atom with the classical partial charge. The corresponding treatment in a standard LCAO-MO solution would require expressions for coupling the classical world to the off-diagonal elements of the Hartree-Fock matrices.

The use of the trace corresponds to optimizing the total energy of an electron density that surrounds a set of atomic nuclei. As such, explicit orthonormal basis functions are not required, although treating exclusion and spin is still necessary. Unlike standard LCAO-MO methods, trace methods do not produce a set of orbitals or a set of orbital energies. Therefore, the trace is not useful for directly estimating electronic transitions nor is it directly useful for producing HOMO or LUMO terms for estimating chemical reactivity. However, trace methods are highly suitable for treating quantum problems that are embedded in a classical problem because the classical perturbation of the trace is readily found.<sup>8</sup> For many large problems the ability to couple the classical approximation to the quantum system is more important than the ability to predict electronic transitions. This is especially true with calculations on biological macromolecules.

The trace method is implemented in the program AMMP.<sup>9,10</sup> The implementation in AMMP uses a subminimal basis with the number of "orbitals" being equal to the number of electrons. This is essentially a minimally parameterized density functional method. An orbital in this implementation means a group of basis functions that cosegregate in an exchange integral. The energy of the system is directly optimized by changing the parameter values of the expansion.

---

## Energy Function

Density functional methods can be readily developed from the standard LCAO-MO approaches by using the trace theorem of linear algebra. The Roothian treatment of the Hartree-Fock equations for the Hamiltonian or total energy of a closed shell system results in the eigensolution of the

following expression<sup>11</sup> for the electron energies:

$$|h_{\text{kinetic}} + h_{e-\text{nuc}} + 2J - K - \lambda e| = 0,$$

where the kinetic energy is

$$h_{\text{kinetic}}(ij) = \left\langle \varphi_i \left| -\frac{1}{2} \nabla^2 \right| \varphi_j \right\rangle,$$

the electron–nuclear energy is

$$h_{e-\text{nuc}}(ij) = \left\langle \varphi_i \left| -\frac{Z_a}{r} \right| \varphi_j \right\rangle,$$

and the electron–electron energy is

$$J(ij) = \left\langle \varphi_i \varphi_j \left| \frac{1}{r} \right| \varphi_k \varphi_l \right\rangle$$

for Coulomb terms and

$$K(ij) = \left\langle \varphi_i \varphi_l \left| \frac{1}{r} \right| \varphi_j \varphi_k \right\rangle$$

for exchange terms.

The total energy is the electron energy plus the internuclear terms,

$$\sum_{a < b} \frac{Z_a Z_b}{r_{ab}},$$

where  $Z_a$  is the atomic number of the nucleuse,  $\lambda$  is an eigenvalue,  $e$  is the identity matrix,  $\varphi$  is a basis element, and  $(ij)$  stands for the  $i, j$  element of the matrix. In all of the equations the summations over  $i, j$  are over electron basis functions, and the sums over  $a, b$  are over atom centers. In the standard treatment the basis is assumed to be orthonormalized, which results in simple expressions for the integrals. The energy is found from the sum of the eigenvalues. Typically more basis functions ( $\varphi$ ) are used than there are electron pairs so that there are more eigenvalues than necessary to find the total energy. Only those eigenvalues that correspond to occupied electrons are summed. The flexibility required to fit the electron density is achieved by using a large number of linear terms in the basis. Only those eigenvalues that correspond to occupied electrons are summed. In a subminimal basis the number of basis functions

are the same as the number of electron pairs, and all the eigenvalues are summed because all the orbitals are occupied. A subminimal basis will only converge if all of the parameters of the density are adjusted for optimal energy. An equivalent way of stating this is that a subminimal basis has the rank of the Slater determinant equal to the number of electron pairs, while in a minimal or superminimal basis the rank of the determinant is much higher. If the degree of the system is the same as the number of electron pairs, then the trace of the matrix is the total energy of the system.

$$H = \sum_{\text{occupied}} \lambda_i = \sum_{\text{all, subminimal}} \lambda_i$$

The trace theorem of linear algebra states that for a Hermitian matrix  $h$ , the sum of the eigenvalues and the trace of the matrix are identical. (The trace theorem applies to the solution of a matrix system of equations and should not be confused with the concept of operators of “trace class.”)

$$\sum_i \lambda_i = \sum_i h_{ii}.$$

The trace theorem follows simply from the result that the trace of a matrix is invariant with respect to any similarity transformation. The eigenvectors of a Hermitian matrix define a similarity transformation that results in a matrix with the eigenvalues on the diagonal.

The apparent simplicity of orthogonalization is misleading. When the orthogonal basis functions are expanded in a linear combination of nonorthogonal basis functions

$$\varphi_k = \sum_i c_{ki} \phi_i,$$

then all of the individual basis integrals still have to be calculated to find the integrals over the orthogonal basis. The overlap integral,

$$\phi_k^* \phi_l = \sum_i \sum_j c_{ki}^* c_{jl} \phi_i^* \phi_j,$$

requires just as many function evaluations in the orthogonalized and nonorthogonalized coordinates.

The terms of the trace may be calculated for nonorthogonal basis elements by the use of the Löwdin formula.<sup>12</sup> The two-centered integrals for an arbitrary operator  $X$  are written as

$$\sum_{i,j} \langle \varphi_i | X | \varphi_j \rangle T_{i,j},$$

where  $T_{i,j}$  is the element of the inverse of the overlap matrix  $\langle \varphi_i | \varphi_j \rangle$ . Four-centered integrals are written as

$$C \sum_{i,j,k,l} \langle \varphi_i \varphi_j | X | \varphi_k \varphi_l \rangle T_{i,l} T_{j,k}$$

where  $C$  is constant, depending on the spin integrals (two for the coulomb integrals, one for the exchange integrals). These highly useful formulae can be readily derived from the analytical expression for an element in the inverse of a matrix in

terms of the matrix's adjoint and determinant

$$A_{ij}^{-1} = \text{adj}(A_{ij}) / \det(A).$$

Single electron operators require evaluation of integrals in the general form of<sup>11</sup>

$$\begin{aligned} & \langle D | H | D' \rangle \\ &= \int \cdots \int (\varphi_1^*(1), \dots, \varphi_{N'}^*(N)) (O(i)) \\ & \times \begin{vmatrix} \varphi_1(1) & \varphi_2(1) & \cdots & \varphi_{N'}(1) \\ \varphi_1(2) & \varphi_2(2) & \cdots & \varphi_{N'}(2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_1(N) & \varphi_2(N) & \cdots & \varphi_{N'}(N) \end{vmatrix} \\ & \times dr_1 \cdots dr_N. \end{aligned}$$

Expanding the matrix and distributing the integrations results in the matrix expression

$$\begin{vmatrix} \int \varphi_1^*(1) \varphi_1(1) & \int \varphi_1^*(1) \varphi_2(1) & \cdots & \int \varphi_1^*(1) \varphi_{N'}(1) \\ \vdots & \vdots & & \vdots \\ \int \varphi_i^*(i) O(i) \varphi_1(i) & \int \varphi_i^*(i) O(i) \varphi_2(i) & \cdots & \int \varphi_i^*(i) O(i) \varphi_{N'}(i) \\ \vdots & \vdots & \ddots & \vdots \\ \int \varphi_N^*(N) \varphi_1(N) & \int \varphi_N^*(N) \varphi_2(N) & \cdots & \int \varphi_N^*(N) \varphi_{N'}(N) \end{vmatrix}.$$

Because the basis functions are not orthogonal, they cannot simply be replaced with zero or one.

However, the cofactor expansion can be used to sum the basis products.

$$\sum_j \int \varphi_i^*(i) O(i) \varphi_j(i) \begin{vmatrix} \int \varphi_1^*(1) \varphi_1(1) & \cdots & 0 & \cdots & \int \varphi_1^*(1) \varphi_{N'}(1) \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots \\ \int \varphi_N^*(N) \varphi_1(N) & \cdots & 0 & \cdots & \int \varphi_N^*(N) \varphi_{N'}(N) \end{vmatrix} = \sum_j \int \varphi_i^*(i) O(i) \varphi_j(i) \text{Adj}(\varphi^* \varphi)$$

Löwdin's formula simply uses the relationship between the adjoint and the inverse [i.e.,  $\text{adj}(A)_{ij} = (A^{-1})_{ij}/\det(A)$ ] to evaluate this determinant.

$$\begin{aligned} & \sum_{i,j} \varphi_i^* O(i) \varphi_j \text{adj}(\varphi^* \varphi) \\ &= \det(\varphi^* \varphi) \sum_{i,j} \varphi_i^* O(i) \varphi_j (\varphi^* \varphi)^{-1}_{i,j}. \end{aligned}$$

The determinant of the overlap matrix is a constant that is removed when the wave function is normalized. The normalized value is simply

$$\sum \int \varphi_i^* O(i) \varphi_j (\varphi^* \varphi)^{-1}_{i,j}.$$

A similar derivation, which starts from the density matrix ( $\varphi^* \varphi$ ), can be applied to the calculation of the two-electron operators required for electron-electron interactions. The Löwdin formulae can be applied to many operators and are used for the electrostatic field and electron density calculations. Although at first glance the use of orthogonal orbitals appears to be simpler, the expansion of the orbitals in terms of a basis set requires evaluation of exactly the same integrals as the Löwdin formulae.<sup>12</sup>

The Hamiltonian with the Löwdin<sup>12</sup> expressions is then

$$\begin{aligned} H = & \sum_{a < b} \frac{Z_a Z_b}{r_{ab}} \\ & + \sum_{i,j} \left( \left\langle \varphi_i \left| -\frac{1}{2} \nabla^2 \right| \varphi_j \right\rangle + \sum_a \left\langle \varphi_i \left| \frac{Z_a}{r} \right| \varphi_j \right\rangle \right) T_{i,j} \\ & + \sum_{ij} \sum_{kl} \left\langle \varphi_i \varphi_j \left| \frac{1}{r} \right| \varphi_k \varphi_l \right\rangle (2T_{ij} T_{kl} - T_{il} T_{jk}). \end{aligned}$$

In standard LCAO-MO theory the extreme value of the Hamiltonian is found by a Picard iteration in the eigenspace of the Hamiltonian. All the  $i, j$  elements of the electronic Hamiltonian are calculated and the lowest energy eigenvectors (the eigenvectors with the lowest eigenvalues) are chosen to represent the wave function for the next iteration. This process converges to a stationary point in the energy that is an energy minimum.

The trace of the Hamiltonian is the sum of the energy eigenvalues by the trace theorem of linear algebra. Therefore, the same energy minimum can be found by direct minimization of the trace of the Hamiltonian. The energy is directly minimized by

adjusting the parameters of the expansion of  $\varphi$ . The trace of the electronic Hamiltonian is simply

$$\begin{aligned} H_{\text{electronic}} &= \sum_i \left( \left\langle \varphi_i \left| -\frac{1}{2} \nabla^2 \right| \varphi_i \right\rangle + \sum_a \left\langle \varphi_i \left| \frac{Z_a}{r} \right| \varphi_{ij} \right\rangle \right) T_{i,i} \\ &+ \sum_{ij} \sum_{kl} \left\langle \varphi_i \varphi_j \left| \frac{1}{r} \right| \varphi_k \varphi_l \right\rangle (2T_{ij} T_{kl} - T_{il} T_{jk}). \end{aligned}$$

The trace of the electronic Hamiltonian combined with the internuclear Hamiltonian,

$$\sum_{a < b} Z_a Z_b / r_{ab}$$

is directly optimized in terms of the atomic positions and basis parameters. Considerable savings in computer time are possible when the integrals that are identical,

$$\begin{aligned} & \langle \varphi_i \cdots | X | \varphi_j \cdots \rangle, \\ & \langle \varphi_j \cdots | X | \varphi_i \cdots \rangle, \end{aligned}$$

are calculated once and then reused in the calculation. The complexity of the calculations can also be reduced when small integrals (corresponding to small elements in the density matrix) are skipped.

The energy expression derived above is a density functional expansion because it operates in terms of density matrices. Other density functional expansions have been derived<sup>13</sup> (for a recent description of the field see Liu and Parr<sup>14</sup>). From the fundamental theory of density functional expansions<sup>13</sup> (an especially clear presentation is in Zaremba<sup>15</sup>) they must be physically equivalent, although the numerical properties of convergence will be different. Standard density functional treatments differ from this expression by expanding the electrons in spin densities rather than pair densities, and they use a different approximation to the exclusion integrals.<sup>14,15</sup> However, for single determinant expansions the use of the trace greatly simplifies the mathematics (see Donnelly<sup>16</sup>). The variational steps in density functional methods are equivalent to the optimization steps required to use this algorithm because both search for a stationary point of the energy that is also an energy minimum.

## Another Approach to Expressions for Hamiltonian

The expressions for the Hamiltonian were derived by converting a standard eigensolution approach on a subminimal basis into a direct optimization approach using the trace of the matrices. They can also be derived from the variational expressions for density functional theory. For example, the electron–nuclear potential in Hohenberg and Kohn's<sup>13</sup> eq. (3) is

$$V \equiv \int v(r) \varphi^*(r) \varphi(r) dr.$$

Inserting a Slater determinant of rank  $N$  for  $\varphi$  results in the familiar determinantal expansion

$$V = \int \begin{vmatrix} \phi_1^*(1) & \cdots & \phi_N^*(1) \\ \vdots & & \vdots \\ \phi_1^*(N) & \cdots & \phi_N^*(N) \end{vmatrix} \times v(r) \begin{vmatrix} \phi_1(1) & \cdots & \phi_N(1) \\ \vdots & & \vdots \\ \phi_1(N) & \cdots & \phi_N(N) \end{vmatrix} dr,$$

where  $dr$  is the integration over all the individual coordinates. Gathering terms results in<sup>11</sup>

$$\begin{aligned} \langle D|H|D' \rangle &= \int \cdots \int (\varphi_1^*(1), \dots, \varphi_N^*(N)) (v(r_{ij})) \\ &\times \begin{vmatrix} \varphi_{1'}(1) & \varphi_{2'}(1) & \cdots & \varphi_{N'}(1) \\ \varphi_{1'}(2) & \varphi_{2'}(2) & \cdots & \varphi_{N'}(2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{1'}(N) & \varphi_{2'}(N) & \cdots & \varphi_{N'}(N) \end{vmatrix} \\ &\times dr_1 \cdots dr_n, \end{aligned}$$

which is the same expression as above after relating the  $v(r)$  as the equivalent operator. Equation (2) in Hohenberg and Kohn<sup>13</sup> can be changed from

$$\frac{1}{2} \int \nabla \varphi^* \nabla \varphi dr$$

to the more familiar

$$\frac{1}{2} \int \varphi^* \nabla^2 \varphi dr$$

by integration by parts, remembering that the wave function goes to zero at infinity. Similar operations can be performed on the electron–electron terms [eq. (4) in Hohenberg and Kohn<sup>13</sup>] to result in the expressions used in this article.

## Classical Trace Integrals

The contribution of the classical parts of the problem to the quantum part is treated by calculating the expected perturbation of the trace of the Hamiltonian. Unlike a whole-matrix LCAO approach, it is only necessary to consider the perturbation to the trace in order to couple the classical and quantum parts of the calculation. The simplicity of this perturbation term is a major reason for using the trace rather than an eigenfunction expansion. Coupling of the larger classical system to the quantum system via the trace or density should hold for other density functional theories as well. The perturbation of the Hamiltonian for both the potential energy and kinetic energy operators must be calculated for the perturbation to be valid. Because the calculation of the perturbation in the kinetic energy does not have a simple expression but it can be treated numerically, the simplification of using the trace is critical to the success of the calculations.

The perturbation of the potential energy operator is given by the 6–12 internuclear distance term with the classical partial charge electrostatic energy with both the nuclei and the electrons. The electrostatic terms for the electrons are simply the electron–nuclear energy integral described above.

$$\begin{aligned} h_{\text{trace}} &= \sum_{\text{nuclei, partial}} \frac{Z_a q_c}{r_{ac}} - \frac{a_a a_c}{r_{ac}^6} + \frac{b_a b_c}{r_{ac}^{12}} \\ &- \sum_{i,j} \langle \varphi_i | \frac{q_c}{r} | \varphi_j \rangle T_{i,j}, \end{aligned}$$

where  $q_c$  is the classical charge and the  $a_x, b_x$  terms are the coefficients of the 6–12 Lennard–Jones potential. The purely classical part of the Hamiltonian is treated by conventional sums over covalent geometry, partial charges, and van der Waals terms.

Although the form of the potential presented above is somewhat heuristic, it can be placed on a more rigorous basis by considering the operations necessary to take one large quantum system and split it into two small systems. The large system

has an overlap matrix of the form

$$\begin{bmatrix} \varphi_1^* \varphi_1 & \varphi_1^* \varphi_2 & \varphi_1^* \varphi_3 & \cdots & \varphi_1^* \varphi_N \\ \varphi_2^* \varphi_1 & \varphi_2^* \varphi_2 & \varphi_2^* \varphi_3 & \cdots & \varphi_2^* \varphi_N \\ \varphi_3^* \varphi_1 & \varphi_3^* \varphi_2 & \varphi_3^* \varphi_3 & \cdots & \varphi_3^* \varphi_N \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \varphi_N^* \varphi_1 & \varphi_N^* \varphi_2 & \varphi_N^* \varphi_3 & \cdots & \varphi_N^* \varphi_N \end{bmatrix}.$$

In order to separate the two subsystems, a similarity transformation is used to force the off-diagonal elements corresponding to the overlap between systems to be zero. Because the systems no longer overlap, it should be possible to solve them independently and construct a perturbation that reflects the coupling between them. The overlap matrix is now

$$\begin{bmatrix} \chi_1^* \chi_1 & \cdots & \chi_1^* \chi_J & & & \\ \vdots & \ddots & \vdots & & & \\ \chi_J^* \chi_1 & \cdots & \chi_J^* \chi_J & & & \\ & & & \chi_{J+1}^* \chi_{J+1} & \cdots & \chi_{J+1}^* \chi_N \\ 0 & & & \vdots & \ddots & \vdots \\ & & & \chi_{J+1}^* \chi_N & \cdots & \chi_N^* \chi_N \end{bmatrix}.$$

The similarity transformation, which is just a coordinate transformation, does not change the trace of this matrix or any matrix involved in the calculation of the Hamiltonian. However, it changes the specific values of the matrix elements and will not generally force the off-diagonal elements of the Hamiltonian to zero. Therefore, it would be necessary to write perturbation equations for at least all the elements in the diagonal block with the standard eigenvector approach to solving the equations. Additionally, the coupling between the systems in the eigenvector approach will depend on the values of the off-diagonal blocks of the Hamiltonian. The results will critically depend on how the system is partitioned. Ideally the off-diagonal blocks of the Hamiltonian will have uniformly small elements that can be safely neglected. Unfortunately, although it may be possible to define heuristics for partitioning the system that usually result in small off-diagonal blocks, the formal definition of a general algorithm for this problem is difficult.

The new coordinate system is a linear combination of the old basis elements; the diagonal elements in the new coordinate system contain terms

that relate the interaction energies of the old systems. The trace of the Hamiltonian in the new coordinate system consists of the energy of the isolated systems added to the interaction energy between the systems. The perturbation assumption is that the newly separated states can be solved individually and an estimate of the interaction energy between the newly separated systems can be applied to correct the energy. Because only the trace is required for the solution, the form of the perturbation is simple because it is not necessary to estimate perturbations for off-diagonal elements. The limit in the use of the trace is that the perturbation be valid. If a classical approximation, like one derived from a molecular mechanics potential, is used then the molecular fragment must be large enough that the atoms of chemical interest and their covalently bonded neighbors are in the fragment. The criteria for the classical approximation are established by the Feynman–Hellmann theorem<sup>17</sup> that shows how to expand the energy for a stationary state in a power series and the connection between these expansions and perturbation theory. The requirement for stationarity means that the atoms of chemical interest must be wholly embedded in the quantum system, but the more distal interactions between the atoms of interest and the rest of the system can be treated by a power series expansion that is represented by a molecular mechanics potential.

The trace of the kinetic energy operator should also be perturbed to maintain the balance between kinetic and potential energy. This is required for the wave function to be stationary. Unlike the potential energy operator, there is no simple expression for the perturbation in the kinetic energy. However, numerical solutions to the wave equation must maintain a balance between the kinetic and potential energies to be valid physical solutions. The balance between the kinetic and potential energy is described by the virial ( $H - T$ ), which is used as a measure of error in the calculations for a closed quantum system. The virial sum ( $H + T$ ) should be identically zero at the solution. The virial sum is violated when the classical potential energy perturbation is used without a correction to the kinetic energy. The correction to the kinetic energy is made by requiring that the virial sum be zero. The approximation that the virial sum be constrained to zero for the open classically perturbed system results in a stable quantum calculation.

## CINDO/INDO Approximations

The use of individual basis elements has an important practical consequence. INDO approximations neglect terms in the Hamiltonian when the expected contribution to the energy is small.<sup>18</sup> This is similar to the local density approximation in density functional theory. Because the Löwdin  $T_{i,j}$  weights are always calculated,<sup>12</sup> they can be used to estimate when the energy integrals are small and thereby apply the INDO approximation at a fine grain. Neglecting terms when  $T_{i,j}$  is less than 0.0001 results in acceptable errors. The use of the INDO approximation is an option in AMMP, and it can be useful for speeding up the initial stages of optimization where the accuracy is not critical.

## Basis Set

A basis set is chosen to be a simple minimal basis set that uses linear combinations of Gaussians to represent the electron density. The basis is a direct extension of the FSGO approach.<sup>6</sup> This results in low code complexity because only one analytical form needs to be integrated for all of the energy integrals. Analytical expressions for the integrals are taken from Boys<sup>19</sup> and Shavitt.<sup>7</sup> The orbital geometry is specified by molecular internal coordinates. For example, a basis element that would represent a  $\sigma$  bond is specified by two atoms and the fractional distance along the interatomic vector. Basis elements used to specify lone pairs are described by three atoms. The cross product of the two bond vectors and the mean of the two bond vectors define the coordinates for the basis center. The  $\pi$  orbitals are represented by pairs of Gaussians with opposite signs. The axis of the  $\pi$  orbital is defined by the cross product of three atoms. Basis elements defined in this way transform with the molecular geometry. Once solved in one configuration, a good approximation to the energy for another configuration can be found without reoptimization of the system.

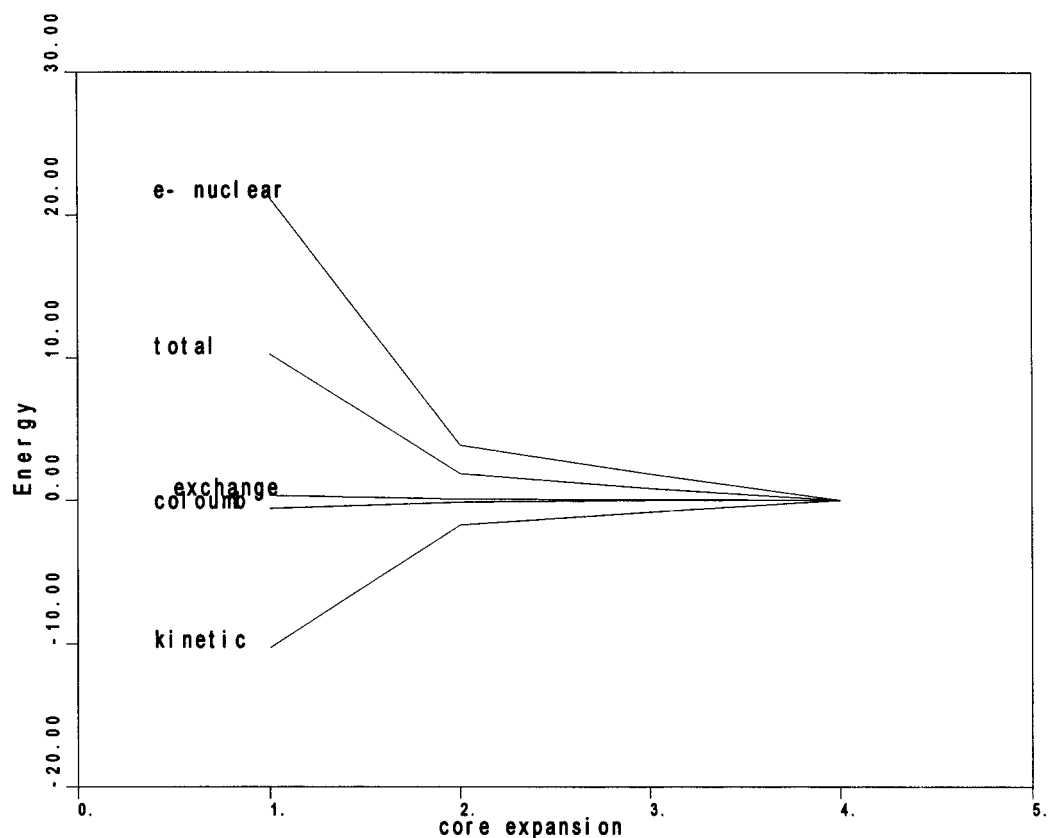
Single Gaussian minimal basis sets or FSGOs poorly reproduce the total energy of a system. Typical errors are on the order of 10–20% of the Hartree–Fock limit. It is a legitimate question whether such a severely minimal basis is of any use. Fortunately, most of this error is due to poor representation of the highly peaked core orbitals

by a single Gaussian. For example, with methane the FSGO energy is  $-33.99$  Hartrees,  $-38.15$  Hartrees with two Gaussians, and  $-39.13$  Hartrees when four Gaussians are used to represent the core  $1s^2$  orbitals ( $r$  C—H =  $1.1126$  Å, H—C—H =  $109.47^\circ$ ). The STO-3G energy is  $-39.718$  Hartrees.<sup>20</sup> The Hartree–Fock limit for methane is approximately  $-40.225$  Hartrees with a correlation error of  $0.295$  Hartrees.<sup>21</sup> Similar convergence is seen with other molecules; the convergence of different terms in the Hamiltonian is shown for ethane in Figure 1. The FSGO energies calculated with the trace method are the same as those presented in the literature<sup>6,22</sup> when the same basis is used. However, many molecular properties do not require that the core orbitals be well represented. For example, in the original FSGO articles<sup>6,22</sup> the errors in covalent geometry were comparable to the errors in a typical molecular mechanics model. Figure 2 shows the lack of dependence of the rotational barrier between eclipsed and staggered conformations on the  $1s^2$  core expansion in ethane, methylamine, methanol, and hydrogen peroxide.<sup>23–26</sup> In all four cases the torsional barriers are estimated to reasonable agreement with the experiment in the location of the minimum energy configuration and the depth of the well. In these calculations the electronic structure was optimized in one of the energy minima (either  $60$  or  $90^\circ$ ). The energy was calculated for the other conformations without further optimization. Tracking with optimization does not alter the results. Figure 3 shows the comparison between the classical molecular mechanics potential and the quantum potential for the single Gaussian core expansion with ethane.

## Specific Basis Definitions

All of the basis elements in this work can consist of a linear sum of up to eight Gaussians with a common center. This simple basis was chosen to avoid numerical difficulties while retaining sufficient flexibility to show that density functional expansion parameterized in terms of internal molecular geometry is transformed with changes in the internal geometry. Obviously, high accuracy results will require better bases. It should be possible to construct similar bases for use with spin densities in standard density functional theory. The way the molecular geometry is coupled to the bases will still be important and will be shaped by the results presented. Multiple copies of this sum





**FIGURE 1.** The convergence of various terms in the Hamiltonian for ethane is shown as a function of the core expansion. The energy difference from the four term expansion for the core is shown.

with different relative locations can be used to represent complicated orbitals. Symmetry can be enforced by grouping or ganging parameters together so that they have identical values. Explicit multicenter basis elements can be made by using either a defined multicenter s-type basis element or an explicit linking command.

#### **SINGLE ATOM CENTERED BASIS ELEMENT**

Single atom centered basis elements correspond to s orbitals in standard bases. The atom that is the atom center is specified. A single center basis element in which the center is displaced by an arbitrary vector was tested but was not found to be useful.

#### **ATOM PAIR CENTERED BASIS ELEMENT**

Atom pair centered elements are very useful for representing  $\pi$  bonds and lobes on  $sp^2$  orbitals such as a carbonyl oxygen. The two atoms are specified and the location of the basis element center is constrained to lie on the line that the

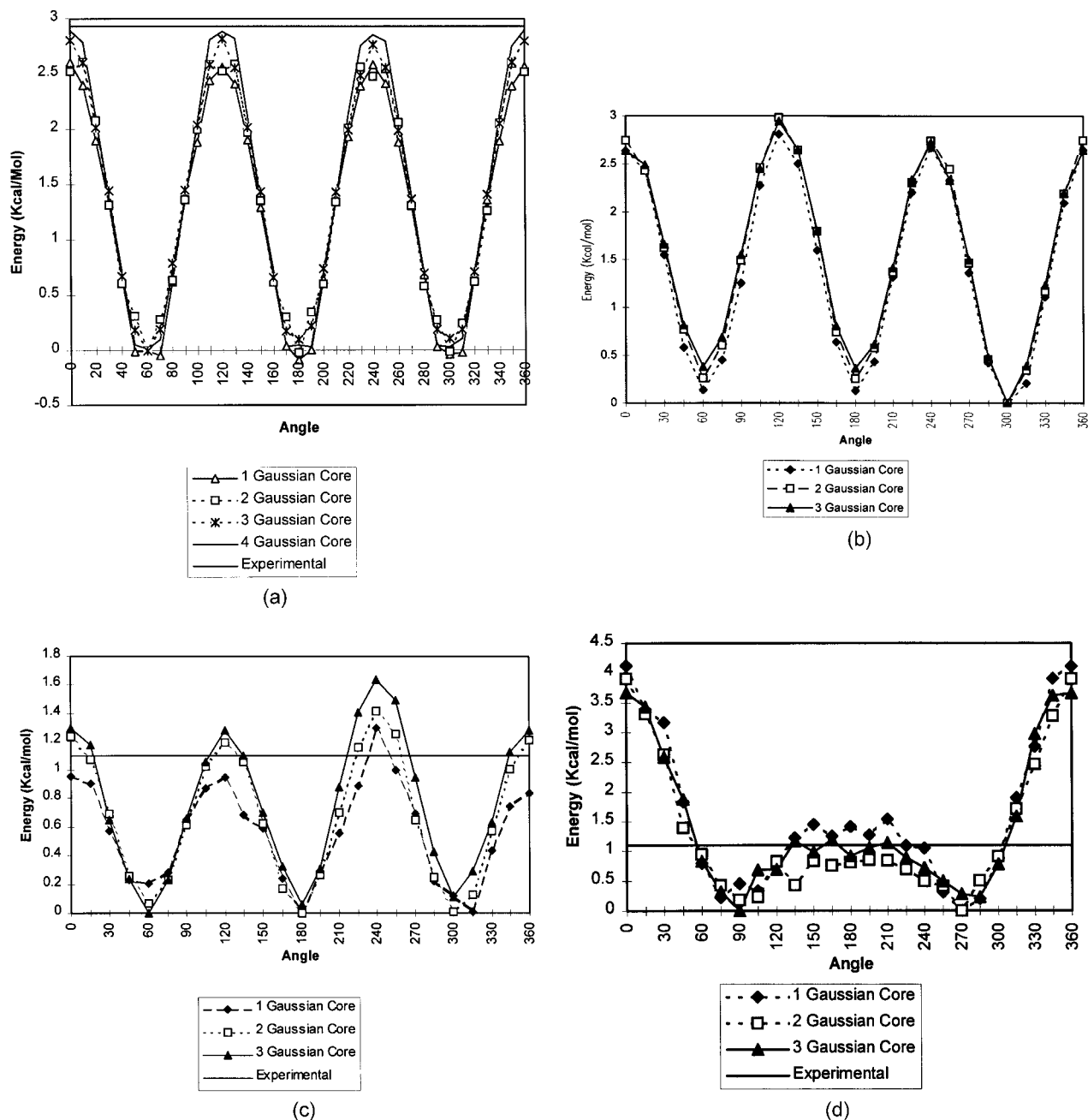
atoms define. The fractional distance along that line from the first atom to the second is used to specify the precise center. (A fractional distance of one would correspond to the center of the second atom.) Dummy atoms or orbital guide atoms can be used with atom pair centered functions to specify directions for orbitals such as the lone pairs in an  $sp^2$  carbonyl oxygen, lone pairs in  $sp^3$  oxygen, and the lone pair in  $sp^3$  nitrogen.

#### **ATOM TRIPLE CENTERED BASIS ELEMENT**

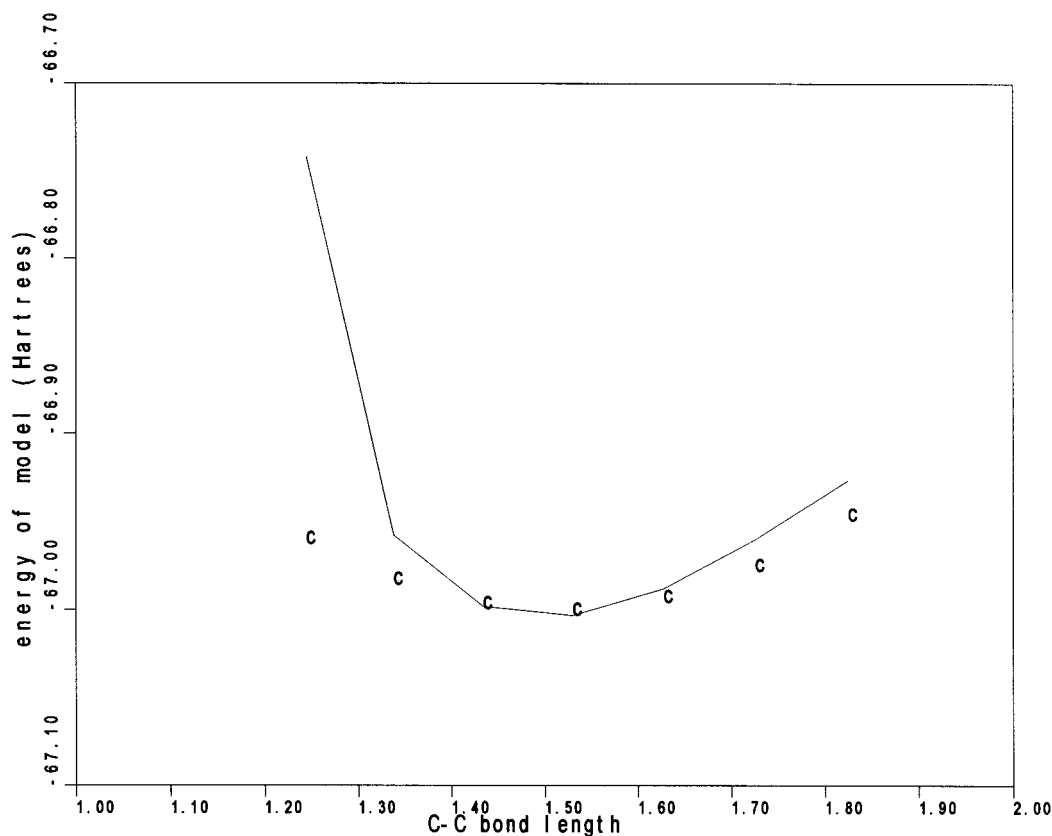
A triple of atoms defines an angle. The bisector and cross product define a unique position for the center of the basis element in the plane of the bisector. The bisector and cross product vectors are normalized and the position is defined by two fractional coordinates.

#### **FOUR ATOM CENTERED BASIS ELEMENT**

Four atoms define a volume. Two forms of four atom basis elements are used. In each case one is



**FIGURE 2.** (a) The difference in the energy of ethane is plotted as a function of the torsion angle. The four curves show the dependence of the energy difference on the expansion of the core orbital. Even though there are large changes in the total energy with each expansion, the energy differences as a function of conformation are nearly identical. The experimental value for the barrier of  $2.93 \pm 0.03$  kcal/mol is shown as a solid line.<sup>23</sup> (b) The torsional energy for methylamine is plotted here for different expansions of the core orbitals. The experimental barrier of 2 kcal/mol is overestimated in this calculation.<sup>24</sup> (c) The torsion energy for methanol is shown for the single, double, and triple Gaussian core expansion. The experimental value of 1.1 kcal/mol is shown.<sup>25</sup> (d) The torsion energy for hydrogen peroxide is shown for different expansions of the core electrons. While the trans barrier of 1.1 kcal/mol is well estimated, the cis barrier of 7 kcal/mol is underestimated.<sup>26</sup> The prediction of a trans barrier shows that the calculations are treating exchange in a fundamentally correct manner.



**FIGURE 3.** The energy as a function of bond length in ethane is shown for a single Gaussian expansion of the core. The classical energy value derived from IR spectra [harmonic with  $K = 349.8 \text{ kcal}/(\text{mol } \text{\AA}^2)$ , origin superimposed on the quantum value] is shown by the points labeled c. Even with a low order expansion for the core electrons, reasonable values are found for deviations around observed bond lengths.

defined to be a center and the normalized cross product of the vectors among the other three is used to define an orientation vector. Single centered basis elements are used to represent lone pairs in trigonal systems like  $\text{NH}_3$ . The positions of the hydrogen atoms define the orientation vector, and the basis element center is defined by a fractional coordinate along the orientation vector from the position of the nitrogen atom. Double centered basis elements are used to represent  $\pi$  orbitals. Here three atoms define an orientation, but two lobes are defined with opposite sign. The fractional coordinate defines the offset of the positive lobe. It is often valuable to use the same triple of atoms as a reference in conjugate planar systems like benzene and avoid problems of arbitrary orientation of  $\pi$  orbitals. Orbitals that are  $\pi$ -like can also be defined along the vector defined by pairs

of atoms. These terms can be useful for defining lone pairs in water and similar systems.

### MULTICENTERED BASIS ELEMENT

Multicentered basis sets can be defined in two ways. An explicit basis element was defined in terms of an arbitrary set of up to six atom centered expansions. However, the alternative approach of defining a higher level command in the command language that would couple arbitrary basis elements together is much more useful. For example, benzene ( $\text{C}_6\text{H}_6$ ) has 42 electrons and requires 21 pairs (21 is not divisible by 6). Therefore, no combination of the single centered orbitals defined above is correct for benzene. Either symmetry or total electron number is violated. However, defining  $\pi$  orbitals for each carbon atom and grouping

them in pairs results in the correct number of basis functions.

### SYMMETRY CONSTRAINTS AMONG BASIS ELEMENTS

Molecules possess considerable symmetry. The symmetry can be used to greatly reduce the difficulty of a calculation by lowering the number of free parameters and to improve the quality of the results by minimizing purely numerical errors. Symmetry is achieved by requiring that sets of parameters vary identically in the calculation. The grouping or ganging together of different parameters is treated as a logical rather than a numerical constraint. Basis elements that are ganged together have the same expansion and the same geometric parameters.

### IMPLEMENTATION NOTES

The program AMMP uses a compilation model of computation. Molecules and molecular properties are compiled into programs that AMMP then executes. In keeping with this philosophy short subprograms were written that set up individual chemical classes of atoms. Most of the recent work with AMMP uses a variant of the UFF potential,<sup>1,27</sup> and the orbital expansions were chosen to have the same logical structure as the atom types in the potential.

### Correctness Check

The calculations depend on the correctness of the evaluation of the integrals. A simple yet rigorous test is to calculate the energies of isolated atoms and ions and compare the values to accurately known ones,<sup>28,29</sup> as well as values from recent density functional studies.<sup>14</sup> Most of the drawbacks of the simple basis set used in this work are irrelevant for these isolated symmetric problems, so the ultimate accuracy of the calculation should be good. The total energy and the exchange energy are presented so that the accuracy of the terms of the Hamiltonian and the accuracy of the total energy can be examined. In each case the results from an optimal multiple Gaussian expansion are shown in Table I.<sup>14,28,29</sup> The errors in the energies are comparable to those in Liu and Parr.<sup>14</sup>

A series of small molecule systems are also included in Table II.<sup>11,20–22,30,31</sup> The H<sub>2</sub> molecule

**TABLE I.**  
**Energy Comparisons.**

Atom	Exchange	Exchange (Exact)	Total	Total (OPM)
He (S)	–1.028	–1.0258	–2.820	–2.862
Li <sup>+</sup>	–1.653	–1.65	–7.211	
Be (S)	–2.674	–2.6658	–14.534	–14.573
B <sup>+</sup>	–3.493	–3.49	–24.187	
C (P)	–5.044	–5.05	–37.442	–37.690
O (P)	–8.186	–8.17	–74.093	–74.814

This table shows how well the energies compare between accurate calculations for total energy and exchange energy<sup>14,27,38</sup> and those found using the trace for isolated atoms and ions. The atoms and ions were chosen so that only electron pair integrals were required for their evaluation.

solves to a total energy of –1.122 with an exchange energy of –0.654 and a bond distance of 0.741 Å with a three Gaussian expansion that compares well with near Hartree–Fock values.<sup>21</sup> LiH has a total energy of –7.948 with an exchange value of –2.15 at a bond length of 1.61 Å (1.595 Å experimentally). Li<sub>2</sub> solves to a total energy of –14.739 with an exchange energy of –3.578 and a bond distance of 2.61 Å (2.66 Å experimentally). These energies are reached only with expansions that are more complicated than simple sums of Gaussians. The 1s terms must be included for each hydrogen atom in the hydrides with at least three Gaussian terms in the expansion, and the bonding electrons must be expanded with at least two Gaussians. The coupling construct described above was used to ensure that the correct number of pairs of electrons was in the expansion.

**TABLE II.**  
**Results for Series of Small Molecules.**

Molecule	Energy	STO-3G	Limit Value
H <sub>2</sub>	–1.122	–1.117	–1.134
LiH	–7.948		–7.986
Li <sub>2</sub>	–14.823		–14.872
BH <sub>3</sub>	–26.202	–26.065	–26.33
CH <sub>4</sub>	–39.799	–39.718	–40.225
NH <sub>3</sub>	–55.400	–55.454	–56.225
H <sub>2</sub> O	–75.395	–74.963	–76.065
H <sub>2</sub> CO	–112.352	–112.353	–113.93

The calculations achieve comparable accuracy to STO-3G<sup>11,20,30</sup> when using a comparable basis. The Hartree–Fock limit values are from Hariharan and Pople,<sup>21</sup> Palke and Lipscomb,<sup>31</sup> Christoffersen,<sup>11</sup> and Frost et al.<sup>22</sup>

Optimization

The ground state is found by minimizing the total energy. Typical initial values for coefficients and exponents are shown in Table III. Finite difference conjugate gradients with inexact line search and the Poliak–Ribeire  $\beta$  is used in AMMP. While the naive implementation of finite differences results in an  $O(N^5)$  complexity calculation [ $N$  calculations of  $O(N^4)$ ], caching of four centered integrals reduces the complexity to  $O(N^4)$  [ $N$  calculations of  $O(N^3)$ ]. This is the same complexity as when using analytical gradients. However, finite differences are less numerically precise than analytical gradients and so analytical gradients will be included in future development. Analytical gradients will also be more important when more complete expansions of the electron density are included. The two centered integrals used in the classical perturbation of the trace are also cached because their calculation can be expensive. Gradients are estimated by using both positive and negative differences with a step size of  $10^{-4}$ . This value was chosen to give reliable gradients for a wide range of conditions. Differences in the inverse of the overlap matrix, which are required by the Löwdin treatment of the Slater determinant,<sup>12</sup> are required as are differences in the basis integrals to generate accurate gradients. Because different parameters of the Hamiltonian have different convergence properties, it is important to allow the independent optimization of different classes of parameters. For example, refining the basis geometry or atomic coordinates before optimizing the expansions of the basis elements generally results in poor quality solutions. The energy is most dependent on the parameters of the expansion of the electron density, less dependent on the basis

geometry, and least dependent on the atomic coordinates. Therefore, an arbitrary set of parameters for the expansion of the electron will not allow the correct refinement of the basis geometry, while an approximate set of parameters for the basis geometry will allow the expansion to be refined. The virial (the difference between the kinetic and total energy) is a highly useful measure of solution quality in addition to the total energy of the system. Unlike the total energy of the system the value for the virial at the minimum energy is known. The scaled virial  $(H - T)/H$  should be identically 2 at the solution, and the virial sum  $H + T$  should be zero. While the optimization of the electron density will converge to low virials without additional terms, including the virial in the energy function often improves the rate of convergence near the solution.

Calculation of Electrostatics and Partial Charges

One major use of quantum calculations is the calculation of molecular electrostatic potential. Classical partial charges for use in molecular mechanics<sup>32</sup> can be determined by fitting to the molecular electrostatic potential. The electrostatic potential is given a modification of the electron–nuclear energy integral

$$\sum_{i,j} \langle \varphi_i | \frac{1}{r} | \varphi_j \rangle T_{i,j},$$

where  $r$  is the distance between the sample point and the integration volume. The electron density is found from

$$\sum_{i,j} \varphi_i \varphi_j T_{i,j},$$

where the basis functions are not integrated, but the value at the sample point is used. The electrostatic potential is then fit by a set of partial charges under the constraint of a total charge. Sample points less than 2.0 Å from an atom center are excluded because the classical point charges are poor at reproducing the field near the atoms. It may be advisable to use a kernel other than  $1/r$  for the electrostatics near the atoms, and the algorithm can be trivially modified to include another kernel that reflects the nonpoint nature of the charge density at small separations. Calculation of partial charges and electrostatic fields requires that

TABLE III. Typical Initial Values for Parameters.

Basis Element	Coefficient	Exponent	Geometry Parameter
C 1s <sup>2</sup>	3.8	9.3	
C, N, O — H lobe	1	−1	0.6
C, N, O — C, N, O lobe	1	−1	0.5
N 1s <sup>2</sup>	4.9	13	
O 1s <sup>2</sup>	6	17.3	

the expansion of the core orbitals use several Gaussians to effectively screen the nuclear charge. Three-dimensional maps of the electron density and the molecular electrostatic potential (both the classical and the quantum) may be produced with AMMP.

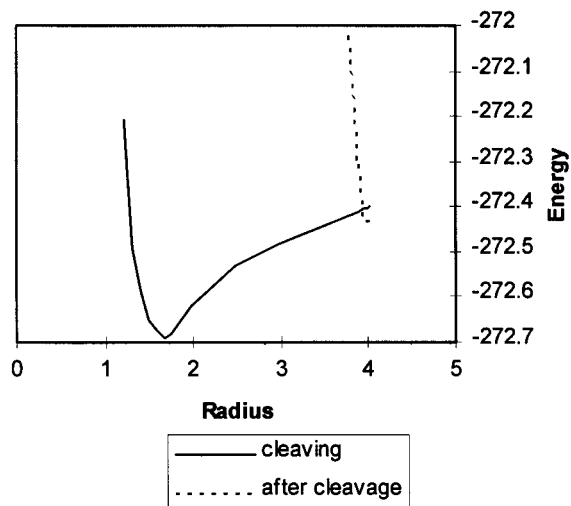
## Results and Discussion

The utility of trace calculations combined with molecular mechanics are illustrated with sample calculations on peptide bond cleavage and on ionizable groups in the complex between HIV protease and the inhibitor Indinavir that are performed to analyze the protonation states. Quantum treatment of different complexes is important for this problem because these calculations involve changes in covalent bond structure that are due to the addition or subtraction of hydrogen atoms on the inhibitor. This will also demonstrate that the trace method is readily embedded in a classical field.

### PEPTIDE BOND CLEAVAGE

A major motivation for this work has been the treatment of transition states and other nonequilibrium structures. The cleavage of the peptide bond is a pharmacologically important example of amide hydrolysis. Often the reaction proceeds via the formation of a tetrahedral intermediate that is structurally analogous to a *gem*-diol amine. Molecular models of complexes of the *gem*-diol amine with aspartyl proteases have shown significant correlation between experimental reaction velocities and molecular mechanics estimates of interaction energies.<sup>1,33a</sup> The main drawback in a molecular mechanics model of a kinetic intermediate is the *ad hoc* nature of the model for the transient species.

In order to understand the molecular geometry of the tetrahedral intermediate a model was made for the *N*-methyl acetamide tetrahedral intermediate. The C—N bond length was varied from 1.22 to 4 Å and the potential energy was calculated for both the tetrahedral and planar geometry around the carbonyl carbon. As shown in Figure 4, when the carbon was tetrahedral the curve of potential versus bond length shows behavior typical of a chemical bond, albeit a strange C—N bond with an equilibrium distance of about 1.6 Å. When the structure is allowed to rearrange to make the



**FIGURE 4.** The reaction profile is shown for hydrolysis of a *gem*-diol amine model of the transition state in peptide bond cleavage. The profile shows a typical bond length profile, albeit with an unusually long bond distance.

O—C—O atoms planar, the C—N bond is no longer evident and the energy of the system shows typical van der Waals repulsion with a minimum at 3.98 Å. The transition state barrier (maximum energy—minimum energy) is approximately 90.3 kcal/mol. More accurate estimates using PM3 and MP2/6-31g\*\*//4-31G calculations on formamide place the transition state of energy of  $\text{NH}_3^+ - \text{CH}(\text{OH})_2$  at 47.1 and 36.8 kcal/mol.<sup>5,34</sup> These results are consistent with the observations of Weber and Harrison<sup>1,33a</sup> that the tetrahedral intermediate is not the transition state of the reaction, but is rather a transient chemical intermediate along the reaction pathway.

### INDINAVIR

AMMP was used to model and rationalize kinetics and inhibition in HIV protease. HIV protease is a dimeric aspartic acid protease whose active site is on the dimer interface. HIV protease inhibitors are a major therapeutic agent in treating AIDS, and understanding viral resistance to these inhibitors is critical. The molecular mechanics interaction energy calculated with AMMP correlates well with estimates of reaction rates<sup>1,33a</sup> and variations in  $K_i$  with changes in the protein.<sup>35</sup> For example, with the neutral inhibitor saquinavir the correlation between  $\ln(K_i)$  and the interaction energy is 0.82 for the wild type and nine resistant mutations. However, treatment of ionizable groups

is always difficult in the molecular mechanics formalism. Indinavir, which is another clinically relevant inhibitor, includes a pterazine ring that is ionizable. It is therefore critical that the correct ionization state be chosen. Performing correlation or regression analysis<sup>33a,b</sup> to compare molecular mechanics and experimental energies without establishing an independent rationale for the choice of ionization state can be problematic. Because Indinavir has four ionization states in the piperazine ring alone, the degrees of freedom could easily be increased fourfold without a corresponding increase in the number of data. Dimethyl piperazine, which also contains a pterazine ring, has a *pK* at 9.66 and 5.20.<sup>36</sup> Therefore, at least the monoprotonated form of the inhibitor will be present under normal reaction conditions (typically pH 5–6.5). A further complication is the close proximity of a pterazine ring nitrogen to the aspartic acids. The protonation environment for the aspartic acids could be asymmetric as a result of the close proximity of the pterazine ring and aspartic acid 25'.

To address these problems in a more rigorous method, quantum mechanics was used to calculate the ground state energy of different dimethyl piperazine ionization states overlaid with HIV protease and the rest of Indinavir. Atomic coordinates were from the crystal structure of the Indinavir complex (pdb1hsg.ent).<sup>37</sup> The calculation embedded in the enzyme uses the 66 electrons of dimethyl piperazine and the 3598 atoms of the protease, ordered solvent, orbital guide atoms, and inhibitor. Because it is a chemical base, the diprotonated form of dimethyl piperazine should be the most stable, but the important question is which of the two possible monoprotonated forms is more stable in the context of the enzyme active site. The energies when embedded in the enzyme are shown in Table IV. The energy of the hn1 (+1) complex is lower than the hn2 (+1) complex. The hn1 complex is the more stable of the two possible (+1)

complexes. Therefore, the only monoprotonated state that needs to be considered is the n1 protonated state. n1 is closer to the active site than n2.

## Conclusions

Moderate accuracy *ab initio* quantum mechanics can be readily integrated into molecular mechanics using density functional techniques. The use of the trace theorem of linear algebra converts an iterative eigenvalue problem (Hartree–Fock-LCAO) into a straightforward optimization problem. The optimization problem is readily solved by classic methods, such as finite difference conjugate gradients, in a reasonable amount of time. The trace method is amenable to use on a small part of a macromolecular problem and will be useful for studying chemical reactions in the context of otherwise intractably large catalysts. It will also be useful for parameterizing classical potentials used in modeling, molecular mechanics, and dynamics.

## References

1. Weber, I. T.; Harrison, R. W. *Protein Eng* 1996, 9, 679–690.
2. Holloway, M. K.; Wai, J. M.; Halgren, T. A.; Fitzgerald, P. M. D.; Vacca, J. P.; Dorsey, B. C.; Levin, R. B.; Thompson, W. J.; Chen, L. J.; deSolms, S. J.; Gaffin, N.; Ghosh, A. K.; Giuliani, E. A.; Graham, S. L.; Guare, J. P.; Hungate, R. W.; Lyle, T. A.; Sanders, W. M.; Tucker, T. J.; Wiggins, M.; Wiscount, C. M.; Woltersdorf, O. W.; Young, S. D.; Darke, P. L.; Zugay, J. A. *J Med Chem* 1995, 38, 305–317.
3. Field, M. J.; Bash, P. A.; Karplus, M. *J Comput Chem* 1990, 11, 700–733.
4. Bernardi, F.; Olivucci, M.; Robb, M. A. *J Am Chem Soc* 1992, 114, 1606–1616.
5. Liu, H.; Müller-Plathe, F.; van Gunsteren, W. F. *J Mol Biol* 1996, 261, 454–469.
6. Frost, A. A.; Rouse, R. A. *J Am Chem Soc* 1968, 90, 1965–1969.
7. Shavitt, I. *Methods Comput Phys* 1963, 2, 1–45.
8. Feynman, R. P. *Statistical Mechanics—A Set of Lectures*; Addison-Wesley: Reading, MA, 1972; p 39–46.
9. Harrison, R. W.; Chatterjee, D.; Weber, I. T. *Proteins Struct Funct Genet* 1995, 23, 463–471.
10. Harrison, R. W. *J Comput Chem* 1993, 14, 1112–1122.
11. Christoffersen, R. E. *Basic Principles and Techniques of Molecular Quantum Mechanics*; Springer-Verlag: New York, 1989; Sections 11-1, 11-2, p 442.
12. Löwdin, P.-O. *J Chem Phys* 1950, 18, 365–375.
13. Hohenberg, P.; Kohn, W. *Phys Rev B* 1964, 3, 864–871.
14. Liu, S.; Parr, R. G. *J Comput Chem* 1999, 20, 2–11.

**TABLE IV.**  
**Energy of Dimethyl Piperazine Ionization States.**

State	Energy (Hartrees)	Energy Difference from +2 State
+ 2	– 296.128	0
+1 (hn1)	– 295.466	0.662
+1 (hn2)	– 294.688	1.440
+0	– 292.471	3.657

15. Zaremba, E. In *Lecture Notes in Physics* 187; Keller, J.; Gázquez, J. L., Eds.; Springer-Verlag: Berlin, 1983; p 167–217.
16. Donnelly, R. A. In *Lecture Notes in Physics* 187; Keller, J.; Gázquez, J. L., Eds.; Springer-Verlag: Berlin, 1983; p 89–125.
17. Stanton, R. E. *J Chem Phys* 1962, 36, 1298–1300.
18. Almöf, J.; Faegri, K., Jr.; Korsell, K. *J Comput Chem* 1982, 3, 385–399.
19. Boys, S. F. *Proc R Soc* 1950, A200, 542–554.
20. Hehre, W. J.; Stewart, R. F.; Pople, J. A. *J Chem Phys* 1969, 51, 2657–2664.
21. Hariharan, P. C.; Pople, J. A. *Theor Chem Acta* 1973, 28, 213.
22. Frost, A. A.; Prentice III, B. H.; Rouse, R. A. *J Am Chem Soc* 1967, 89, 3064–3065.
23. Weiss, S.; Leroi, G. E. *J Chem Phys* 1968, 48, 962.
24. Lide, D. R., Jr. *J Chem Phys* 1957, 27, 343.
25. Ivash, E. V.; Dennison, D. M. *J Chem Phys* 1953, 21, 1804.
26. Hunt, R. H.; Leacock, R. A.; Peters, C. W.; Hecht, K. T. *J Chem Phys* 1965, 42, 1931.
27. Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard III, W. A.; Skiff, W. M. *J Am Chem Soc* 1992, 114, 10024–10035.
28. Engel, E.; Vosko, S. H. *Phys Rev A* 1993, 47, 2800–2811.
29. Li, Y.; Krieger, J. B.; Iafrate, G. *J Phys Rev A* 1993, 47, 165–181.
30. Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Macmillan: New York, 1982; p 192.
31. Palke, W.; Lipscomb, W. N. *J Am Chem Soc* 1966, 88, 2384–2393.
32. Besler, B. H.; Merz, K. M., Jr.; Kollman, P. A. *J Comput Chem* 1990, 11, 431–439.
33. (a) Weber, I. T.; Harrison, R. W. *Protein Sci* 1997, 6, 2365–2374; (b) Weber, I. T.; Harrison, R. W. In *3D QSAR in Drug Design. Recent Advances*; Kubinyi, H.; Folkers, G.; Martin, Y., Eds.; ESCOM Science Publishers: Leiden, The Netherlands, 1997; Vol. 2, p 115–127.
34. Krug, J. P.; Popelier, P. L. A.; Bader, R. W. F. *J Phys Chem* 1992, 96, 7604–7616.
35. Weber, I. T.; Harrison, R. W. *Protein Eng* 1999.
36. Lide, D. R., Ed. *CRC Handbook of Chemistry and Physics*, 72nd ed.; CRC Press: Boca Raton, FL, 1991; p 8–37.
37. Chen, Z.; Li, Y.; Chen, E.; Hall, D. L.; Darke, P. L.; Culbertson, C.; Shafer, J. A.; Kuo, L. C. *J Biol Chem* 1994, 269, 26344.